DESCRIPTION

IMAGE RECOGNITION APPARATUS AND IMAGE RECOGNITION PROGRAM

5   Technical Field

The present invention relates to an image recognition apparatus capable of advantageously recognizing even the substance of an image included in sports-related contents such as a image material in an uncompleted state for

10   broadcasting and a sports program being telecasted or as recorded in a recording medium, the substance of such an image having been conventionally difficult to recognize.

Background Art

15   Among conventional methods of recognizing the substance of an image, such as "successful passing shot" and "successful smash", contained in a scene of a sports picture of, for example, tennis, there are known methods of recognizing the substance of such an image by manually

20   inputting a "successful passing shot" section, a "successful smash" section and a like section of picture information one by one, or by extracting positions of respective of a ball, players and court lines and totally judging a change with time in spatial correlations among the extracted positions with use

25   of a computer.

The method of image recognition based on manual

input, however, involves a problem of increased labor costs and a problem of heavy burden on the operators which arises when the contents processing takes a long time, though the substance of an image can be reliably recognized. On the

5   other hand, the method of automatic image recognition with a computer has a difficulty in judging whether the ball is on line or out particularly when the ball falls to the ground adjacent a line.

10   Disclosure of Invention

In order to solve the foregoing problems the present invention provides the following means.

That is, the present invention provides an image recognition apparatus for recognizing movements of players

15   matched against each other between domains partitioned with such an obstacle as net in a sport match or game from contents including a television program being telecasted to show the sport match or game, an image material in an uncompleted state for broadcasting and contents recorded in such a recording

20   medium as a VTR, the image recognition apparatus comprising: a score information obtaining section configured to obtain score information indicative of scores of the respective players which vary as the sport match or game proceeds; a play event information obtaining section configured to obtain play event

25   information indicative of a characteristic movement of each of the players from picture information included in the contents,

the picture information containing images of respective of the obstacle and the players, the score information displayed on a screen, and like images; and an image substance recognizing section configured to make a comparison between a score

5   information item obtained immediately before a point in time of generation of the play event information and a score information item obtained immediately after the point in time and make reference to a result brought by the play event information, thereby recognizing a substance of an image

10   provided by the play event information.

This image recognition apparatus is capable of recognizing a specific image substance by exactly grasping the description of the play event information since the apparatus can reliably judge which of the players' respective scores

15   relates to a play event information item from score information items obtained immediately before and after the point in time of the generation of the play event information item.

In a specific embodiment of the present invention

20   for extracting the score information from the contents, the score information obtaining section is configured to obtain the score information from at least one of the picture information included in the contents, sound information including commentary voice of a commentator, and data

25   information transmitted as multiplexed on radio waves during broadcasting.

To extract the characteristic movement of each player from the contents, it is preferable that the image recognition apparatus further comprises: a domain element extracting section configured to extract from the picture information facility information including information on the obstacle, information on the domains and information on boundary lines between the domains and an area outside the domains, player's position information indicative of a player's position, and instrument information on an instrument moving between the domains to serve as an object of score count in the sport match or game; rule information storage section configured to store rule information on the sport match or game; and basic movement storage section configured to store basic movement information on players' characteristic movements generalized in a sport of concern, wherein the play event information obtaining section includes a play event information determining section configured to determine a play event information item on a play event characteristic of each of the players included in the picture information as the play event information based on domain elements extracted from the picture information, the rule information, and the basic movement information stored in the basic movement information storage section.

To extract players' characteristic movements more efficiently, it is desirable that the player's position information be position information indicative of a domain

containing each of the players and the instrument constantly
held and used by the player.

In a specific embodiment of the present invention
for extracting the player's position information from the
picture information, the domain element extracting section is
configured to extract the player's position information from
the picture information based on the facility information
extracted by the domain element extracting section. In a
specific embodiment of the present invention for extracting
the instrument information from the picture information, the
domain element extracting section is configured to extract the
instrument information including information on a ball from
the picture information based on the facility information and
the player's position information extracted by the domain
element extracting section.

For the image recognition apparatus to be capable of
recognizing images including an image of a more complicated
movement of each player, it is desirable that the play event
information obtaining section include a play event index
information output section configured to output plural play
event information items determined by the play event
information determining section as arranged in a time sequence.
It is more desirable that the play event index information
output section be configured to output the play event
information items together with instrument information items
as arranged in a time sequence.

To extract a contents element related to a sport of concern suitably, it is desirable that the facility information, the player's position information, the instrument information and the rule information be based on knowledge about a sport as a subject for image extraction.

Brief Description of Drawings

Fig. 1 is a diagram showing the device configuration of an image recognition apparatus according to an embodiment of the present invention.

Fig. 2 is a function block diagram of the embodiment.

Fig. 3 is a diagram illustrating a court model for use in extracting court lines from picture information according to the embodiment.

Fig. 4 is a diagram illustrating a net model for use in extracting net lines from picture information according to the embodiment.

Fig. 5 is a diagram illustrating the court lines and net lines extracted from picture information according to the embodiment.

Fig. 6 is an illustration of a player's domain detected according to the embodiment.

Fig. 7 is an illustration of a ball domain detected according to the embodiment.

Fig. 8 is an illustration of a trace of a ball position.

Fig. 9 is an illustration of a manner of storage by a basic movement information storage section of the embodiment.

Fig. 10 is an illustration of a manner of storage by a rule information storage section of the embodiment.

5　　　Fig. 11 is an illustration of a manner of identifying player's basic movements according to the embodiment.

Fig. 12 is a time chart in which a play event information item on each player and a ball position are

10　connected to a score information item according to the embodiment.

Fig. 13 is a flowchart showing a process of image recognition from picture information according to the embodiment.

15

Best Mode for Carrying Out the Invention

Hereinafter, one embodiment of the present invention will be described with reference to the drawings.

Fig. 1 is a diagram showing the device configuration

20　of an image recognition apparatus according to an embodiment of the present invention. Fig. 2 is a function block diagram of the embodiment.

The image recognition apparatus according to this embodiment is configured to recognize characteristic movements

25　of players in a sport match or game from sports contents including a television program being telecasted on a

television receiver TV or being reproduced by a

recording/reproducing device such as a VTR, and such contents

as recorded in a recording medium. As shown in Fig. 1, the

image recognition apparatus includes, as major components

5   thereof, an input-output interface 11 connected to the

television receiver TV and the recording/reproducing device

such as a VTR, an external storage device 12 and internal

memory 13, such as HDD or the like, for storing data, programs

and the like, a CPU 14 configured to operate according to a

10   program stored in the external storage device 12 or the like

to cause the apparatus to function as image recognition

apparatus 1, a user interface 15 comprising a keyboard and a

mouse for receiving user information about the user, and a

like component. The "contents", as used herein, is meant to

15   include pictures including images of movements of players, a

shot taken at such an angle to view a court from obliquely

above along the length of the court and a close-up shot of a

judge or a spectator, and sound including voice of a

commentator and the like. In this embodiment, reference is

20   made to a tennis program as an example of the "contents".

In a functional aspect, the image recognition

apparatus 1 has functions as a score information obtaining

section 101, a domain element extracting section 102, a basic

movement information storage section 103, a rule information

25   storage section 104, a play event information obtaining

section 107 comprising a play event information determining

section 105 and a play event index information output section

106, an image substance recognizing section 108, and a like

section, as shown in Fig. 2, which functions are fulfilled by

the operations of the CPU 14 and the like.

5   These sections will be described in detail.

The score information obtaining section 101 is

configured to automatically obtain score information on scores

of respective players being matched against each other which

vary as the match proceeds by image analysis on picture

10   information displayed on a television receiver. The score

information obtaining section 101 is formed by utilizing the

input-output interface 11 and the like. While the score

information obtaining section 101 is configured to obtain the

score information automatically in this embodiment, the score

15   information may be obtained in any desired manner, for example,

by manual input.

The domain element extracting section 102 is

configured to extract from the picture information facility

information including information on such an obstacle as net,

20   information on a court as partitioned domains, and information

on court lines as boundary lines between the domains and an

area outside the domains, player's position information

indicative of the position of each player, and instrument

information on an instrument moving between half-courts to

25   serve as an object of score count in the sport match. In this

embodiment, the facility information to be extracted comprises

information on the court lines and information on the net

lines; the player's position information to be extracted

comprises position information on each of players 1 and 2

matched against each other; and the instrument information to

5   be extracted comprises information on a tennis ball

(hereinafter will be referred to as "ball"). The facility

information, player's position information and instrument

information extracted by the domain element extracting section

will be generally referred to as domain elements.

10   More specifically, in extracting the facility

information, information on the court lines and information on

the net lines are extracted in this order from the picture

information by reference to a court model specifying court

characteristic points Pc_1, ..., Pc_{14} (hereinafter will be

15   generally referred to as "Pc") as representative points on the

court lines and court lines Lc_1, ..., Lc_9 (hereinafter will be

generally referred to as "Lc") as shown in Fig. 3 and a net

model specifying net characteristic points Pn_1, ..., Pn_3

(hereinafter will be generally referred to as "Pn") as

20   representative points on the net lines and net lines Ln_1 and

Ln_2 (hereinafter will be generally referred to as "Ln") as

shown in Fig. 4.

        First, the court lines are extracted from the

picture information by detecting the court characteristic

25   points. More specifically, at a point in time t=0, initial

characteristic points Pc(0) are given as inputs; court lines

Lc(0) determined by the characteristic points Pc(0) are
transformed into a Hough plane; and then a detection window
Wc(0) having dimensions W_{th} and W_{ro} is provided about each peak
point on the Hough plane. At a point in time t=t, first, a

5   binary image B(t) of an original image and an area around
court lines Lc(t-1) are ANDed to generate a binary image Bc(t)
comprising only the area around the court (hereinafter will be
referred to as "court line binary image"). Subsequently, the
process steps of: subjecting this binary image to Hough

10   transformation line by line; performing peak detection within
the range limited by each detection window Wc(t-1); updating
the court characteristic points Pc(t); subjecting court lines
Lc(t) to Hough transformation again; and updating detection
windows Wc(t), are performed to extract the court lines from

15   the picture image. If a certain court characteristic point is
positioned out of the screen due to panning or the like,
update is achieved by estimating the position of the point
outside the screen based on connecting knowledge on the
assumption that a court characteristic point $Pc_i(t)$ (i=9, 10,

20   12, 13, or 10, 11, 13, 14) indicating the center of the court
is constantly displayed on the screen. For the same reason,
some of initial characteristic points may be omitted. The
"connecting knowledge" is knowledge defined based on such
knowledge used in doing a sport of concern that connecting

25   court characteristic points $Pc_i(t)$ (i=9, 10, 12, 13) for
example with each other on a central area of the court allows

a zone that can have a meaning to be defined on the court model.

Subsequently, the net lines are extracted from the picture information by the following process steps: at a point in time t=0, initial characteristic points $Pn(0)$ are given as inputs; a net line $Ln(0)$ and a detection window $Wn(0)$ are provided for each line in the same manner as with the court lines; at a point in time t=t, an image $Bn(t)=B(t)-Bc(t)$, which is a binary image formed by removing the court line binary image from the binary image of the original image, is generated as a net line binary image; this binary image is subjected to Hough transformation; peak detection is performed within each detection window; and the characteristic points $Pn(t)$ are updated.

In this way the court lines and the net lines can be extracted as shown in Fig. 5.

In turn, the player's position information is extracted by specifying a domain in which overlapping is maximum in binary images formed by removing the court lines and the net lines from the picture image.

More specifically, at a point in time t=t, differences from images that are forwardly and backwardly apart from an image of concern by s frames are found to generate binary images $B_1(t)$ and $B_2(t)$ using appropriate threshold values. Here, $B_1(t)=BIN(I(t)-I(t-s))$, and $B_2(t)=BIN(I(t+s)-I(t))$, wherein BIN is a function making the

parenthesized factor binary. Based on a binary image $B_{diff}(t)$

resulting from an AND operation on these two difference images

and a binary image $B_{label}(t)$ in which those points on an image

$I(t)$ at a point in time $t=t$ which are included in a color

cluster corresponding to a predetermined representative color

of, for example, a players' uniform are each defined as 1, the

court lines and the net lines are erased. Further, a domain

from which a portion overlapping the player's domain is

considered to have been removed is compensated through

expansion/compression processing. The two images thus

obtained are ORed to give a binary image $B(t)$ as shown in Fig.

6. A connected domain within the binary image $B(t)$ thus

obtained is labeled and the thus labeled domain is observed

throughout several frames to avoid influence of noise. Such a

domain in the area covering the court and the area therearound

is determined as a player's initial position if the domain has

an area larger than a predetermined value. Of such domains

each having an area larger than the predetermined value at the

point in time $t=t$, those domains each of which is located

adjacent a player's domain at a point in time $t=t-1$ and has

the smallest difference in area from the latter player's

domain is judged as a player's domain p, thereby providing

player's position information.

By switching between a detection mode and a trace

mode in accordance with the distance from the player's

position given by the player's position information thus

extracted, the ball is extracted.

More specifically, the detection mode is a mode for detecting all ball prospect positions each having an area smaller than a predetermined template $T_b(x,y)$ in a domain

5　around each player within an image I'_B from which the players' positions P have been erased at the point in time t with use of the template $T_b(x,y)$ provided with a ball size of $b_x \times b_y$, as shown in Fig. 7. Likewise, ball prospects at points in time t=t+1, t+2, ... are detected and series of ball prospects Ba

10　which are detected to be radially consecutive from about a player's position are chosen and the number of such series of ball prospects Ba is reduced by selection to find a single series of ball prospects Ba. The finally selected series of ball prospects Ba can be specified as a ball trajectory BW in

15　a time segment of concern. The template $T_b(x,y)$ is a kind of tool provided for extracting the ball from the picture information. In this embodiment the size of the ball to be displayed as expanded or compressed is provisionally established as $b_x \times b_y$ and a periphery slightly expanded

20　outwardly from $b_x \times b_y$ is established as the template.

The trace mode is a mode for tracing the ball trajectory BW by template matching with the template $T_b(x,y)$. In this mode, tracing is conducted using as the center of presumption a position obtained by adding an amount of move

25　detected last time directly to a current frame on the assumption that ball trajectory BW in a very short period of

time can be considered to be substantially straight. When the distance between a player's domain and the position of a ball prospect Ba becomes smaller than a certain threshold value, the trace mode is switched to the detection mode. If not, the

5   trace mode operation is repeatedly conducted.

In this way, the ball trajectory BW in a desired time segment can be obtained as shown in Fig. 8. Note that the ball trajectory BW is superimposed on picture information obtained at a desired point in time in Fig. 8 for convenience

10 in showing the ball trajectory BW.

The basic movement information storage section 103 is configured to store basic movement information on players' characteristic movements generalized in a sport of concern and is provided in a predetermined area of the external storage

15 device 12 or internal memory 13. In this embodiment the basic movement information includes five attribute values as connected with respective basic movement information IDs identifying unique values in the present system, as shown in Fig. 9. The five attribute values are: "stay" indicative of a

20 staying movement of a player on the spot, "move" indicative of a move of a player, "forehand_swing" indicative of a forehand swing motion, "backhand_swing" indicative of a backhand swing motion, and "overhead_swing" indicative of an overhead swing motion.

25 The rule information storage section 104 is configured to store rule information required for a sport of

concern and movement index rule information on movements of players doing the sport and is provided in a predetermined area of the external storage device 12 or internal memory 13. More specifically, as shown in Fig. 10, the movement index

5   rule information includes, for example, a movement index rule information item indicative of a player's movement index "service" defined by the description that both players do "stay" in respective zones outside the court at the same point in time and then either of the players does "overhead_swing"

10   in the zone outside the court based on the basic movement information, and a movement index rule information item indicative of a player's movement index "dash to the net" defined by the description that a player hitting the ball is in a zone other than the net court at the time of hitting and

15   then the player is in the net court at the time of next hitting.

The play event information obtaining section 107 comprises the play event information determining section 105 and the play event index information output section 106 and is

20   configured to obtain play event information items as arranged in a time sequence, each of which is indicative of a player's characteristic movement from picture information included in the contents, the picture information containing information on the obstacle and the players, score information displayed

25   on the screen and like information.

More specifically, the play event information

determining section 105 is configured to determine play event
information item characteristic of each player provided by the
picture information based on the domain elements extracted
from the picture information by the domain element extracting
5   section 102, the rule information and movement index rule
information, and the basic movement information stored in the
basic movement information storage section 103.

Further more specifically, the players' basic
movements, such as "forehand_swing", "backhand_swing" and
10   "overhead_swing", are identified by totally judging the
positional relativity between each player's position and the
ball position and the positional relation between each
player's position and each of the court lines and net lines at
the time of ball hitting. For example, as shown in Fig. 11,
15   if the ball is above an identification line extending through
an upper portion of a rectangle circumscribing a player
hitting the ball at the time of hitting, the movement of the
player is identified as "overhead_swing", while if the ball is
on the foreside or backside with respect to the center of
20   gravity of the player, the movement of the player is
identified as "forehand_swing" or "backhand_swing". The
identification line is established to extend through an upper
portion of a player's domain determined by a fixed proportion
to the length of the player's circumscribing rectangle along
25   the y-axis.

The play event index information output section 106

is configured to output to the play event information

obtaining section 107 a time chart in which plural play event

information items determined by the play event information

determining section 105 and ball position information items

5   are arranged in a time sequence.

More specifically, the time chart describes

information items on the relation between the ball and each of

players 1 and 2 as arranged in a time sequence, as shown in

Fig. 12. For example, the arrow corresponding to

10 "overhead_swing" of the player 1 indicates that during a time

segment of concern the player 1 does "overhead_swing" and the

ball describes a trajectory BW comprising a j-i number of

representative points (p_{i+1}, ..., p_i) on the court.

The image substance recognizing section 108 is

15 configured to make a comparison between score information

items obtained immediately before and after a point in time of

generation of the play event information obtained by the play

event information obtaining section 107 and make reference to

a result brought by the play event information, thereby

20 recognizing the substance of an image provided the play event

information.

More specifically, points in time T1, T2 and T3

immediately following the generation of the play event

information will be described with reference to Fig. 12.

25   First, the player 1 does "overhead_swing"

immediately before the point in time T1 and then does "move"

immediately after the point in time T1. The player 2, on the other hand, does "move" immediately before the point in time T1 and then does "backhand_swing" immediately after the point in time T1. The ball is "hit" about the point in time T1,

5 then does "bounce" and, thereafter, is "hit" again. Since there is no change in the scores of the respective players 1 and 2 with time passing from immediately before to immediately after the point in time T1, the image substance recognizing section 108 recognizes the play event information item on the

10 play done between the players 1 and 2 immediately before the point in time T1 as not a characteristic movement associated with the scores.

Subsequently, the player 1 does "move" immediately before the point in time T2 and then does "forehand_swing"

15 immediately after the point in time T2. The player 2, on the other hand, does "backhand_swing" and, thereafter, does "move" immediately after the point in time T2. The ball is "hit" about the point in time T2. Since there is no change in the scores of the respective players 1 and 2 with time passing

20 from immediately before to immediately after the point in time T2, the image substance recognizing section 108 recognizes the play event information item on the play done between the players 1 and 2 immediately before the point in time T2 as not a characteristic movement associated with the scores.

25 Subsequently, the player 1 does "forehand_swing" immediately before the point in time T3 and then does "stay"

immediately after the point in time T3. The player 2, on the other hand, does "move" and, thereafter, does "stay" immediately after the point in time T3. The ball is "hit" and then does "bounce" immediately before the point in time T3 and,

5   thereafter, becomes "undetected" immediately after the point in time T3. Since the score of the player 1 changes from 0 to 15 with time passing from immediately before to immediately after the point in time T3, the image substance recognizing section 108 recognizes the fact that the ball "hit" by

10   "forehand_swing" of the player 1 does "bounce" on the court of the player 2 and then is not hit back by the player 2, thus giving a point to the player 1. By reference to the basic movement information storage section 103 and the rule information storage section 104 as to a series of such play

15   event information items, the image substance recognizing section 108 recognizes the series of play event information items as indicating a "successful passing".

    Next, the operation of the image recognition apparatus according to this embodiment will be described with

20   reference to the flowchart at Fig. 13.

    Initially, court lines and net lines are extracted from picture information (step S101). Player's position information is extracted from the picture information by specifying a domain in which overlapping is maximum in binary

25   images formed by removing the court lines and the net lines from the picture information (step S102). Based on the

player's position information thus extracted, a ball is
extracted from the picture information (step S103). Play
event information items on a player of concern are determined
based on the domain elements extracted, the rule information,

5   movement index information and basic movement information
(step S104) and then outputted together with ball position
information items as arranged in a time sequence (step S105).
The image provided by the picture information is recognized by
giving a meaning to each play event information item based on

10   a change in score information before and after the generation
of the play event information item (step S106).

Since it can be reliably judged which of the
players' respective scores relates to a play event information
item of concern from score information items obtained

15   immediately before and after the point in time of the
generation of the play event information item as described
above, a relatively inexpensive image recognition apparatus
can be provided which is capable of recognizing a specific
substance of an image by exactly grasping the description of

20   the play event information.

In this embodiment, a tennis program is used as an
exemplary one of the contents, while facility information as a
domain element to be extracted from the picture information on
the tennis program includes information on court lines and

25   information on net lines. It is, however, needless to say
that if the contents change from the tennis program to another

sport program, the facility information to be extracted

changes also. Similarly, the player's position information

and the instrument information also change.

While this embodiment is configured to arrange play

5   event information items in a time sequence and then obtain

score information items before and after the generation of

each play event information item, it is possible to obtain

score information items before and after the point in time of

the generation of each play event information item, for

10   example, without arranging play event information items in a

time sequence. Further, the contents of the play event

information to be obtained and the process for obtaining the

play event information may be embodied variously.

This embodiment is configured to recognize

15   characteristic movements of players playing in a sport match

from sports contents including a television program being

telecasted on a television receiver TV or being reproduced by

a recording/reproducing device such as a VTR, and such

contents as recorded in a recording medium. However, media

20   through which contents as a subject for image recognition are

provided are not limited to those used in this embodiment.

For example, it is possible to recognize characteristic

movements of players playing in a sport match or game from

image materials which have been just taken from the sport

25   match or game at a stadium and hence are in an uncompleted

state for broadcasting or from archived picture information on

Internet.

While this embodiment is configured to extract a ball from picture information using the predetermined template $T_b(x,y)$ having a ball size of $B_x \times B_y$, the ball may be extracted

5   without using the template.

The specific features of other sections or parts are not limited to this embodiment but may be modified variously without departing from the concept of the present invention.

10   Industrial Applicability

According to the present invention having been described above, it can be reliably judged which of players' respective scores relates to a play event information item from score information items obtained immediately before and

15   after the point in time of the generation of the play event information item. Thus, the present invention can provide a relatively inexpensive image recognition apparatus which is capable of recognizing a specific substance of an image by exactly grasping the description of the play event information

20   item.

With such an image recognition apparatus, it is possible to extract a desired portion from a program being telecasted or contents being reproduced by a recording/reproducing device as well as to access archived

25   picture information on Internet or the like for extraction of a desired portion from the picture.